**////Title: Investigating the Reliability of Firearm Examinations**

**////Standfirst:**

In 2009, the US National Research Council Committee on Identifying the Needs of the Forensic Sciences Community published a report highlighting the need to estimate the validity of expert opinions in forensic disciplines. These include the opinions of firearm examiners, who are trained to identify firearms and other weapon-related evidence during criminal investigations. Dr Susan Vanderplas at the University of Nebraska Lincoln has recently introduced a new unifying approach for accurately calculating the error associated with firearm analyses. Her method could be applied in forensic laboratories worldwide, to improve the reliability of forensic evidence in criminal investigations.

**////Main text:**

During criminal proceedings, lawyers often rely on the testimonies of forensic experts specialised in different disciplines, who can perform detailed analyses and offer opinions about specific cases. These experts include firearm examiners, who are trained to examine firearms, cartridges, bullets, and other weapon-related evidence at crime scenes.

Just over a decade ago, the National Research Council Committee on Identifying the Needs of the Forensic Sciences Community published a report that emphasised the pressing need to estimate how often forensic experts make mistakes when investigating a crime scene, also known as their error rates.

Accurately estimating the error rates of firearm examiners and other forensic experts could improve the reliability of forensic procedures, making juries and judges aware of how reliable an expert’s testimony is, while also allowing unreliable testimonies to be excluded from court. This could reduce the likelihood of innocent people being wrongfully convicted and of criminals avoiding conviction due to unreliable evidence.

…

Most previous studies have found that expert firearm examinations had remarkably low error rates. Despite the apparently promising findings gathered by these past studies, very few have highlighted the impact of so-called ‘inconclusive’ analyses and results on the reported error estimates.

To understand inconclusive analysis, consider a case where a firearm examiner is asked to determine whether a bullet at a crime scene was fired from a specific weapon found at a suspect’s home. She must decide whether this evidence is a clear identification, an elimination, or an inconclusive prediction.

An identification means that the examiner identifies a match, an elimination means that she does not, and an inconclusive result means that the available evidence was not substantial enough for her to make a confident prediction.

…

Dr Susan Vanderplas, a researcher who specialises in forensic statistics at the University of Nebraska Lincoln, has recently revisited previous studies that estimated the error rates of firearm examiners, specifically focusing on their treatment of inconclusive results.

Her study was carried out in collaboration with Heike Hofmann and Alicia Carriquiry, two professors and statisticians at Iowa State University. Their paper specifically analysed ‘black-box studies’, which assess examiners’ accuracy rates in the same way as objective laboratory tests: based on whether the conclusions are correct.

…

When approaching classification problems, which entail the grouping of data or items into specific categories, statisticians would generally estimate error rates by calculating the accuracy of predictions. In the case of firearm examination results, however, things are a bit more complicated.

The treatment of inconclusive results when calculating firearm examiner error rates can significantly impact a study’s outcome. If an inconclusive result is considered to be a decision that conveys with certainty the absence of conclusive data, then several inconclusive results would not translate into higher error rates. Contrarily, if they are perceived as incorrect results, then a high number of them would increase error rates.

…

Inconclusive results can be treated in three different ways when estimating the error rates of firearm examinations. They can either be excluded from error calculations, included as correct results, or included as incorrect results.

The way in which researchers treat inconclusive predictions, therefore, can have a huge impact on how the analyses are performed and on the final estimated error rates. As past studies considered these decisions differently, comparing them is not always easy.

In their paper, Vanderplas, Hofmann and Carriquiry clearly outline how classification methods are usually evaluated. These methods are generally assessed based on their sensitivity, their ability to detect whether two objects are derived from a common source, and their specificity, which describes their ability to confidently detect when two items do not match and come from different sources.

A different way of evaluating firearm classification methods involves assessing the probability that two pieces of evidence come from the same source and the probability that they come from different sources.

…

Taking all these considerations into account, Vanderplas, Hofmann and Carriquiry examined the results of ten of the most iconic past studies assessing the error rates of firearms examinations. To do this, they devised a framework that considers all the challenges and elements discussed in their paper.

Overall, the researchers found that there were stark differences in the rates of inconclusive results in different studies and in the ways these results were treated in analyses. They also found that many estimations were unreliable or impossible to calculate, as they were based on problematic experimental procedures or poorly designed experiments.

In the studies considered by the researchers, firearms examiners were more likely to make conclusive decisions when looking at evidence from the same source. When looking at evidence from different sources, examiners were more likely to render inconclusive decisions. This may indicate a process bias or reflect examiner training, but it means that firearms examiners are more willing to evaluate evidence that helps the prosecution and are less willing to decisively evaluate evidence that helps the defence.

…

Vanderplas, Hofmann, and Carriquiry concluded that many existing error rate estimates for firearm examinations are unreliable. They emphasise the need for new studies that are based on large, representative samples of examiners who are asked to make several sets of single-comparison predictions. In addition, the authors suggest that error rates reported from these types of studies are not necessarily the most useful information for juries. Instead, juries should be presented with probabilities that reflect the information known about the case.

At the end of their paper, the researchers also outline elements that would make studies in this area more reliable. For instance, they suggest that more ‘open-set’ studies should be performed, where some of the pieces of evidence presented to an examiner may originate from different sources. Such open-set studies remove any assumption that an examiner might have about whether a comparison is guaranteed to match a given weapon or bullet, for example. The researchers also recommend the random allocation of different test kits to participants, and the inclusion of more challenging samples for examinations.

In the future, the team’s suggestions and proposed experimental approach could be used to estimate the error rates of firearm examiners more reliably in forensic laboratories worldwide. This would give judges and juries a better picture of how accurate the testimonies of firearm examiners are, potentially helping to reduce wrongful convictions and ensuring the sentencing of dangerous criminals.

…

This SciPod is a summary of the paper ‘Treatment of inconclusives in the AFTE range of conclusions’*, Law, Probability and* Risk, 2020, 19, 317–364. [doi.org/10.1093/lpr/mgab002](https://doi.org/10.1093/lpr/mgab002)

For further information, you can connect with Dr Susan Vanderplas at susan.vanderplas@unl.edu